A PROJECT REPORT
ON

FACIAL EXPRESSION RECOGNITION USING LOCAL DIRECTIONAL TERNARY PATTERN (LDTP)

Submitted in partial fulfillment of requirements

For the Project II (IT-452)

B. Tech in INFORMATION TECHNOLOGY

By

CH. TEJASRINIVAS (Y19IT016)
K. HARSHAVARDHAN (Y19IT062)
A. SHYAMSUNDAR (Y19IT001)

NOVEMBER-2022

R.V.R & J.C. COLLEGE OF ENGINEERING (Autonomous)
NAAC A+ Grade, NBA Accredited (Approved by A.I.C.T.E)
(AFFILIATED TO ACHARYA NAGARJUNA UNIVERSITY)
Chandramoulipuram: :Chowdavaram
GUNTUR - 522019

R.V.R & J.C COLLEGE OF ENGINEERING (AUTONOMOUS)

DEPARTMENT OF INFORMATION TECHNOLOGY

BONAFIDE CERTIFICATE

This is to certify that this Project work titled **"FACIAL EXPRESSION RECOGNITION USING LOCAL DIRECTIONAL TERNARY PATTERN"** is the bonafide work of **CH. TEJASRINIVAS (Y19IT016), K. HARSHAVARDHAN (Y19IT062), A. SHYAMSUNDAR(Y19IT001),** who have carried out the work under my supervision, and submitted in partial fulfillment for the award of the degree, **B.TECH.** during the year **2022-2023.**

Smt. N. Neelima
Assistant Professor

Dr. A. Srikrishna
Prof. & HOD, Dept of IT

ACKNOWLEDGEMENT

The successful completion of any task would be incomplete without a proper suggestion, guidance, and environment. Combination of these three factors acts like backbone to our Project Work **“Facial Expression Recognition using Local Directional Ternary Pattern”**

We would like to express our gratitude to the Management of **R.V.R & J.C COLLEGE OF ENGINEERING** for providing us with a pleasant environment and excellent lab facility.

We regard our sincere thanks to our Principal, **Dr. K. Ravindra** for providing support and stimulating environment.

We are greatly indebted to **Dr. A. Srikrishna**, Professor and Head of the Department Information Technology, for her valuable suggestion during the course period.

We would like to express our special thanks of gratitude to our guide **Smt. N. Neelima** who helped us in doing the Project successfully.

We would like to thank our Project in-charges **Smt. G. Swetha,** who gave us the opportunity to do this work.

CH. TEJASRINIVAS (Y19IT016)
K. HARSHAVARDHAN (Y19IT062)
A. SHYAMSUNDAR (Y19IT001)

ABSTRACT

The automatic recognition of emotion has been an important field in computer vision. One of the key techniques for recognizing emotions automatically is facial expression recognition, which detects and analyses human emotions from facial images. Facial expressions can be represented by appearance changes on the face. Consequently, describing them exactly is the key issue in facial expression recognition for detecting emotions. This project proposes a new face descriptor, Local Directional Ternary Pattern (LDTP), for facial expression recognition. Motivated by the high edge responses in the boundaries of the emotion-related facial features, an edge directional patterns in a face image are generated, while avoiding smooth regions (meaningless for expression recognition), by using the magnitude of the edge response. Experiments are carried out on facial expression on CK+ dataset. Performance analysis is performed and results obtained are good.

CONTENTS

List of figures

List of tables

CHAPTER 1

INTRODUCTION

Facial expression recognition software is a technology that uses biometric markers to detect emotions in human faces. More precisely, this technology is a sentiment analysis tool and can automatically detect the six basic or universal expressions: happiness, sadness, anger, surprise, fear, and disgust. Expressions on the face are a vital mode of communication in humans as well as animals. Human behavior, and psychological traits, are all easily studied using facial expressions. It is also widely used in medical treatments and therapies.

Facial expression recognition system is a computer-based technology and therefore, it uses algorithms to instantaneously detect faces, code facial expressions, and recognize emotional states. It does this by analyzing faces in images or video through computer-powered cameras embedded in laptops, mobile phones, and digital signage systems, or cameras that are mounted onto computer screens. The expression recognition systems consist of four major steps which produce an efficient result. The first stage is face detection i.e., the region of interest (ROI) is detected from the image so that computation must be performed only where the required data resides and it increases the computational speed. Along with it, normalization is done which convert data image into a normalized value according to the requirement of application. Next step is feature extraction which extracts the distinct features and irrelevant features are eliminated in feature selection process. Final step of facial expression recognition is classification where the expressions are classified into six basic emotions.

1.1 APPLICATIONS

1. Facial expression recognition has wide range of applications in the field of Human computer interactions, Augment reality, virtual reality, Affective computing, and Advanced driver assistance systems (ADSSS).

2. Market Research: Companies have traditionally done market research by conducting surveys to find out about what consumers want and need. This method however, assumes that the preferences stated are correct and reflect future actions. But this is not always the case. Another popular approach in market research is to employ behavioral methods where user's reactions are observed, while interacting with a brand or a product. Although effective, such techniques can quickly become very labor intensive as the sample size increases. In such circumstances, facial expression recognition technology can save the day by allowing companies to conduct market research and measure moment-by-moment facial expressions of emotions automatically, making it easy aggregate the results.

3. Video Game Testing: Facial expression recognition can also be used in the video game testing phase. In this phase, usually a focus group of users is asked to play a game for a given amount of time and their behavior and emotions are monitored. By using facial expression recognition, game developers can gain insights and draw conclusions about the emotions experienced during game play and incorporate that feedback in the making of the final product. Facial expression analysis is a practical means of going beyond the typical survey approach. It is a way of appreciating what the user is experiencing, all while getting feedback. When feedback is taken in this format, it becomes genuinely non-intrusive when it comes to user experience.

4. The practical application of the detected facial expressions in affective computing are tested using a setup to control various applications that help in improving driver assistance systems, such as speed control and mood lighting.

1.2 DRAWBACKS OF EXISTING SYSTEMS

Many methods like Local Binary Pattern (LBP), Local Directional Pattern (LDP), and Facial Expression Recognition Using PCA and Texture-Based LDN Descriptor were previously used for extracting facial features for facial expression recognition but every method has its own disadvantages apart from having their advantages.

The usage of Local Binary Pattern (LBP) as a Facial Descriptor leads to the limitations like the consideration of Smooth regions for the evaluation besides, they do not contribute anything for facial description and LBP uses larger local regions for facial feature extraction which leads to increase in robustness to errors.

Facial Expression Recognition Using PCA and Texture-Based Local Directional Number (LDN) Descriptor leads to the limitation like as some of the irrelevant features are also extracted as a part of feature extraction which reduces the efficiency of the LDN.

In the usage of Local Directional Pattern (LDP) as a facial descriptor, it uses the smooth regions for evaluation of facial expression recognition which reduces the efficiency of the usage of LDP as a facial descriptor.

1.3 DATA SET FOR FACIAL EXPRESSION RECOGNITION

The dataset used in the project is the CK+ dataset which comprises 100 university students in age 18 to 30 years, of which 65% were female, 15% were African-American, and 3% were Asian or Latino. Subjects were instructed to perform a series of 23 facial displays, six of which were based on descriptions of prototypic emotions (i.e., anger, contempt, disgust, fear, happy, sad, and surprise). Image sequences from neutral to target display were digitized into 48x48 pixel arrays. Three images of every person with the same expression are consisted in this CK+ dataset. All the images present in this dataset were grey-level images of the type Portable Network Graphis (PNG).

1.4 LITERATURE SURVEY

1.4.1 Facial Expression Recognition Using Local Directional Pattern (LDP)

A robust face descriptor is an essential component of a good facial expression recognition system. Analyse the performance of a new feature descriptor, Local Directional Pattern (LDP), for the representation of facial expressions. LDP features are obtained by computing the edge response values in all eight directions at each pixel position and then a code is generated according to the relative magnitude's strength. Thus, each expression is represented as a distribution of LDP codes. Different machine learning techniques are compared using Cohn-Kanade facial expression database for classification. Extensive experiments explicate the superiority of the proposed LDP-based descriptor over other existing well-known descriptors.

1.4.2 Facial Expression Recognition Using PCA and Texture-Based LDN Descriptor

The most expressive way in which humans can display their emotion is through facial expression. The development of automated system that performs this task is very difficult. Texture-based LDN descriptor along with PCA which helps to recognize facial expression in efficient manner. In order to obtain directional information, using two masks i.e. Kirsch and Robinson mask and compare their efficiency.

1.4.3 Face Description with Local Binary Patterns

Efficient facial image representation based on local binary pattern (LBP) texture features. The face image is divided into several regions from which the LBP feature distributions are extracted and concatenated into an enhanced feature vector to be used as a face descriptor. The performance of the proposed method is assessed in the face recognition problem under different challenges.

1.4.4 Graphical Representation for Heterogeneous Face Recognition

Heterogeneous face recognition (HFR) refers to matching face images acquired from different sources (*i.e.*, different sensors or different wavelengths) for identification. HFR plays an important role in both biometrics research and industry. Despite promising progresses achieved in recent years, HFR is still a challenging problem due to the difficulty to represent two heterogeneous images in a homogeneous manner. Existing HFR methods either represent an image ignoring the spatial information, or rely on a transformation procedure which complicates the recognition task. Considering these problems, a novel graphical representation based HFR method (G-HFR) is used. Markov networks are employed to represent heterogeneous image patches separately, which takes the spatial compatibility between neighbouring image patches into consideration. A coupled representation similarity metric (CRSM) is designed to measure the similarity between obtained graphical representations.

1.4.5 Shared Representation Learning for Heterogenous Face Recognition

Heterogenous face recognition is still a challenging problem. The main difficulties are owing to the complex relationship between heterogenous face image spaces. The heterogeneity is always tightly coupled with other variations, which makes the relationship of heterogenous face images highly nonlinear. Many excellent methods have been proposed to model the nonlinear relationship, but they apt to overfit to the training set, due to limited samples. Inspired by the unsupervised algorithms in deep learning, a novel framework for heterogeneous face recognition is used. First extract Gabor features at some localized facial points, and then use Restricted Boltzmann Machines (RBMs) to learn a shared representation locally to remove the heterogeneity around each facial point. Finally, the shared representations of local RBMs are connected and processed by PCA. Near infrared (NIR) to visible (VIS) face recognition problem and two databases are selected to evaluate the performance.

1.5 OBJECTIVES OF WORK

The project aims to recognize the facial expression effectively by using Local Directional Ternary Patterns (LDTP). For experimental validation, we use the well-known CK+ dataset which is collected from Kaggle. Later, how results prove to be prominent when compared to some of the known facial expression recognition techniques will be seen.

1.6 SCOPE OF THE WORK

First chapter, a brief introduction about the facial expression recognition and in the different domains in which they are used efficiently and various methods are discussed under the literature survey. Chapter 2, describes the existing methods and how they can be implemented. Chapter 3, describes how the features are extracted from a given facial image using Local Directional Ternary Patterns (LDTP) for facial expression recognition and trained using classification technique K- Nearest Neighbor. Chapter 4, contains the information of performance evaluation and comparison with other methods. The conclusion is given in Chapter 5.

CHAPTER 2

EXISTING METHODS FOR FACIAL EXPRESSION RECOGNITION

2.1 FACE DESCRIPTION WITH LOCAL BINARY PATTERNS

2.1.1 INTRODUCTION

Local binary patterns (LBP) are a type of visual descriptor used for classification in computer vision. The LBP operator was originally designed for texture description. This operator assigns a label to every pixel of an image by thresholding the 3 X 3-neighborhood of each pixel with the centre pixel value and considering the result as a binary number.

The following are the few advantages of LBP

1. Its invariance to monotonic Gray-level changes and computational efficiency, make it suitable for demanding image analysis tasks.
2. Robustness to challenges such as pose and illumination changes.

2.1.2 METHODOLOGY

This method comprises of two steps which are computation of LBP & Face description with LBP.

2.1.2.1 COMPUTATION OF LBP FOR A GIVEN FACIAL IMAGE

i. The figure 2.1 gives the overall view of LBP operator.

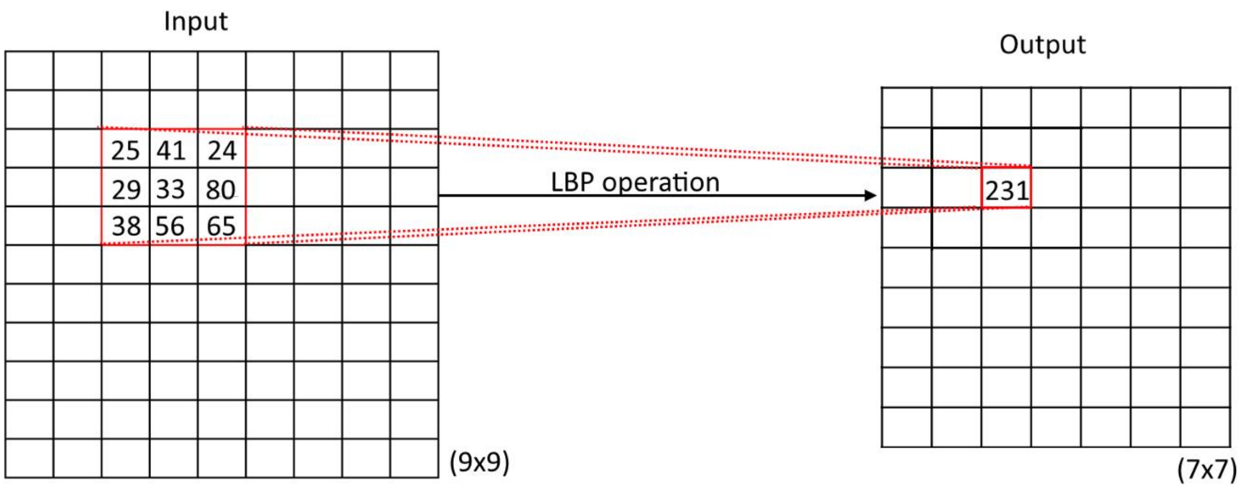


Figure 2.1. Overall view of LBP Operator

ii. Divide the examined window into cells (e.g., 16x16 pixels for each cell).

iii. For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e., clockwise, or counter-clockwise.

iv. Where the center pixel's value is greater than the neighbor's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience).

The Figure 2.2 shows computation of LBP for a given image

10	12	18
7	9	6
9	2	4

If neighbor pixel

≥ **9**	= 1
< **9**	= 0

1	1	1
0		0
1	0	0

Binary Number Generated

1	1	1	0	0	0	1	0

Decimal Number = **226**

Figure 2.2. Local Binary Pattern Code Generation for an Image

A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not. In the computation of the LBP histogram, uniform patterns are used so that the histogram has a separate bin for every uniform pattern and all nonuniform patterns is assigned to a single bin.

2.1.2.2. FACE DESCRIPTION WITH LBP

The procedure consists of using the texture descriptor to build several local descriptions of the face and combining them into a global description. The facial image is divided into local regions and texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face. See the figure for an example of a facial image divided into rectangular regions. The basic histogram is computed for every region that further extended to spatially enhanced histogram which encodes both the appearance and the spatial relations of facial regions. In the spatially enhanced histogram, we effectively have a description of the face on three different levels of locality which are Pixel-Level, Regional Level, Global Level. Each element in the enhanced histogram corresponds to a certain small area of the face.

2.1.3 RESULTS & DISCUSSIONS

The system uses the FERET face images. The FERET database consists of a total of 14,051 Gray-scale images representing 1,199 individuals. The images contain variations in lighting, facial expressions, pose angle, etc. Here only frontal faces are considered. These facial images can be divided into five sets as follows: *fa, fb, fc, dup I, dup II.*

Where *fa* set, used as a gallery set, contains frontal images of 1,196 people.

fb set (1,195 images), The subjects were asked for an alternative facial expression than in the *fa* photograph.

fc set (194 images), The photos were taken under different lighting conditions.

dup I set (722 images), The photos were taken later in time.

dup II set (234 images), This is a subset of the *dup I* set containing those images that were taken at least a year after the corresponding gallery image.

2.1.3.1 COMPARING LOCAL BINARY PATTERNS TO OTHER LOCAL DESCRIPTORS

The comparison of LBP to three other texture descriptors, namely the Gray-level difference histogram, homogeneous texture descriptor, and an improved version of the Texton histogram. The results show that the tested methods work well with the easiest *fb* probe set, which means that they are robust with respect to variations of facial expressions, whereas the results with the *fc* probe set show that other methods than LBP do not survive changes in illumination. The LBP and texton give the best results in the *dup I* and *dup II* test sets.

Table 1 : The recognition rates obtained using different texture descriptors

Method	*fb*	*fc*	*dup I*	*dup II*	lower	mean	upper
Difference Histogram	0.87	0.12	0.39	0.25	0.58	0.63	0.68
Homogenous texture	0.86	0.04	0.37	0.21	0.58	0.62	0.68
Texton Histogram	0.97	0.28	0.59	0.42	0.71	0.76	0.80
LBP (non-weighted)	0.93	0.51	0.61	0.50	0.71	0.76	0.81

The main explanation for the better performance of the local binary pattern operator over other texture descriptors is its tolerance to monotonic grey-scale changes. Additional advantages are the computational efficiency of the LBP operator and that no grey-scale normalization is needed prior to applying the LBP operator to the face image.

The LBP face recognition method calculates histograms over the local regions so a small change in the position of the face relative to the grid causes changes in the labels only on the borders of the local regions. Therefore, it can be expected that the LBP method is not sensitive to small changes in the face localization and that using larger local regions increases the robustness to errors. The effect of localization errors to recognition rate of the LBP method compared to PCA Mah-Cosine was systematically tested as follows: The training images for PCA and gallery (*fa*) images were normalized to size 128 X 128 using provided eye coordinates. The fb set was used as probes. The probes were also normalized to size 128 X 128 but a random vector (Δ X, Δ Y) was added to the face location, where ΔX and ΔY are uncorrelated and normally distributed with mean 0 and standard deviation σ. Ten experiments were conducted with each probe totalling 11,950 queries for each σ tested value.

Table 2: Recognition rates of the LBP & comparison algorithms

Method	***fb***	***fc***	***dup I***	***dup II***	**lower**	**mean**	**upper**
LBP, weighted	0.97	0.79	0.66	0.64	0.76	0.81	0.85
LBP, non-weighted	0.93	0.51	0.61	0.50	0.71	0.76	0.81
PCA, Mah-Cosine	0.85	0.65	0.44	0.22	0.66	0.72	0.78
Bayesian, MAP	0.82	0.37	0.52	0.32	0.67	0.72	0.78
EBGM_Optimal	0.90	0.42	0.46	0.24	0.61	0.66	0.71

The final recognition results fare shown in Table 2 and the rank curves are plotted in Figure. 2.3 LBP yields clearly higher recognition rates than the control algorithms in all the FERET test sets and in the statistical test. The results on the *fc* and *dup II* sets show that specially with weighting, the LBP-based description is robust to challenges caused by lighting changes or aging of the subjects but further research is still needed to achieve even better performance.

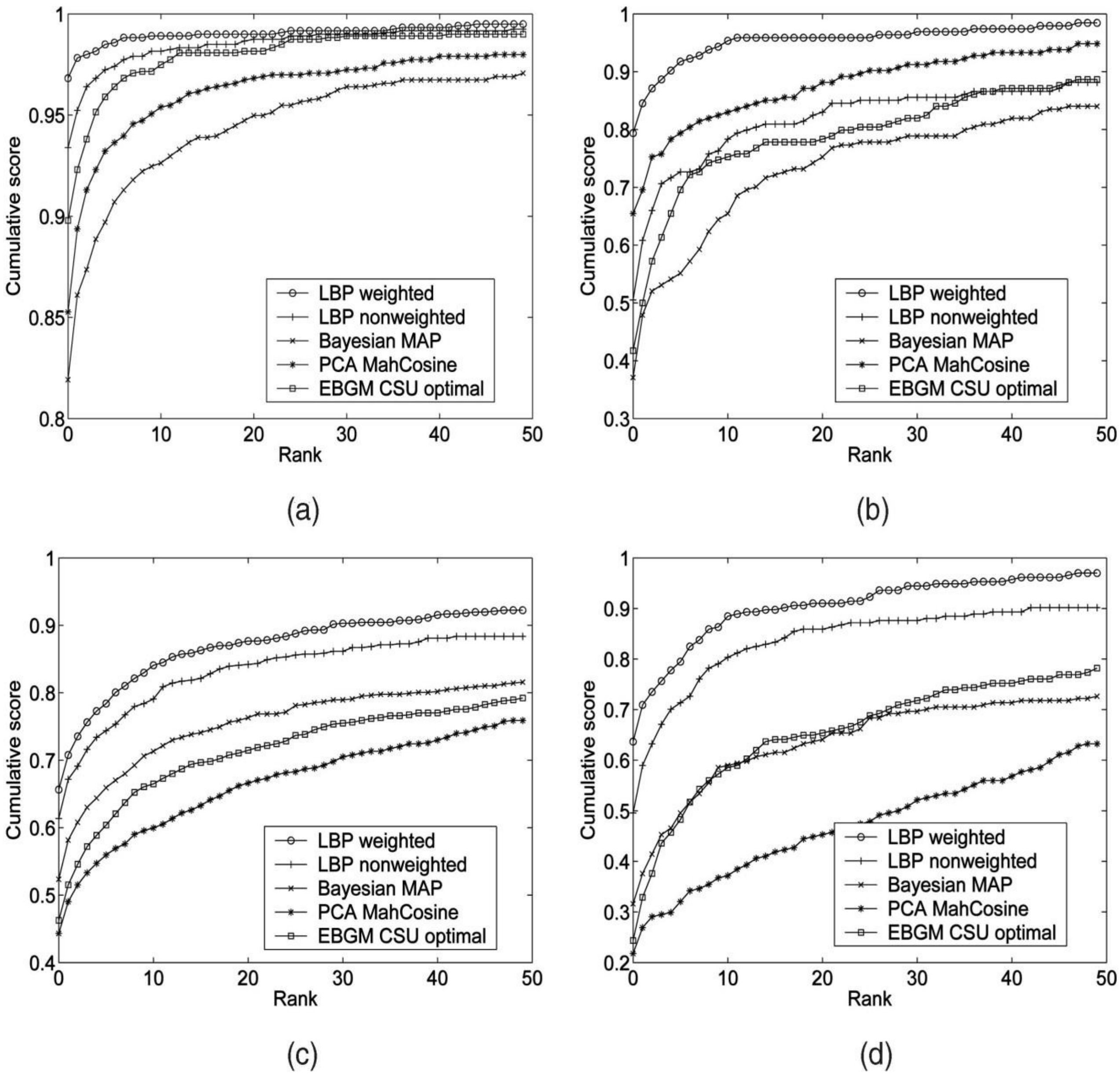


Figure 2.3. The cumulative scores of the LBP and control algorithms on the (a) *fb*, (b) *fc*, (c) *dup I*, and (d) *dup II* probe sets.

2.2 FACIAL EXPRESSION RECOGNITION USING PCA & TEXTURE-BASED LDN DESCRIPTOR

2.2.1 INTRODUCTION

Local Directional Number (LDN) is a face descriptor, which evaluates the directional information of a local neighborhood to encode its structure. Mainly, LDN is a texture descriptor and it encodes the facial texture of structural information and intensity variations.

Principle Component Analysis (PCA) is the most used global descriptor which is used to perform feature selection by eliminating irrelevant features.

In this method, local texture-based descriptors are used for encoding the directional information of the face textures. The main drawback is that some of the irrelevant features are also extracted which reduces the efficiency. To overcome this weakness, pre-processed the input image from the CK+ dataset was by detecting only the face region with the help of the Viola-Jones Algorithm and then evaluated the performance while using Kirsch and Robinson mask. Finally, In this scheme reduction of the dimensionality of the LDN features using PCA to extract relevant features is done.

2.2.2 METHODOLOGY

The directional number pattern is a local feature descriptor for face analysis. It encodes directional information of face textures. Each pixel of the input image is represented as 6 bits of binary code. It helps to distinguish among similar structural patterns that are having different intensity transitions. The main phases that are used to recognize facial expressions in this method include:

Preprocessing Phase: In this phase, we have used face images from the CK+ dataset. In order to obtain the region of interest, i.e., only the face region and to eliminate background Viola-Jones algorithm is used and then crops the face region. This extracted image is then used for further processing thus reducing the computation time.

LDN Code Generation Phase: For generating the LDN code compute the edge response of the neighborhood using a compass mask as it helps to gather information in the 8 directions. In this method, the usage of two compass masks namely Kirsch and Robinson masks are used to extract the facial features. Kirsch mask helps to reveal the structure of the face and it is robust against noise. The edge magnitude is equal to the maximum value found by the convolution of each mask with the image. So, the edge direction is defined by the mask that produces the maximum magnitude. To generate LDN code, analyze the edge response of each mask which represents significant edges in its respective direction. Identify the highest positive and negative values. LDN is the concatenation of minimum and maximum directions. The following Figure 2.4 is the Kirsch mask rotated 45° apart to obtain edge response in 8 different directions.

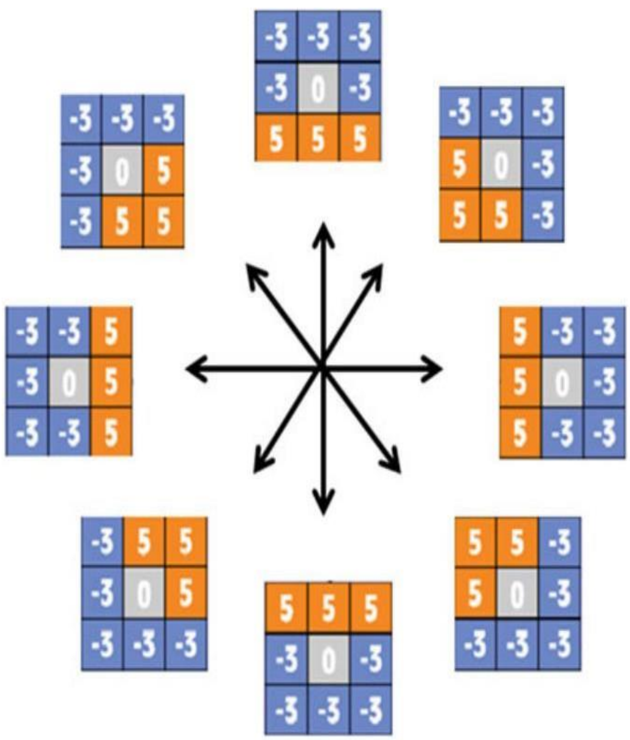


Figure 2.4. Kirsch Mask

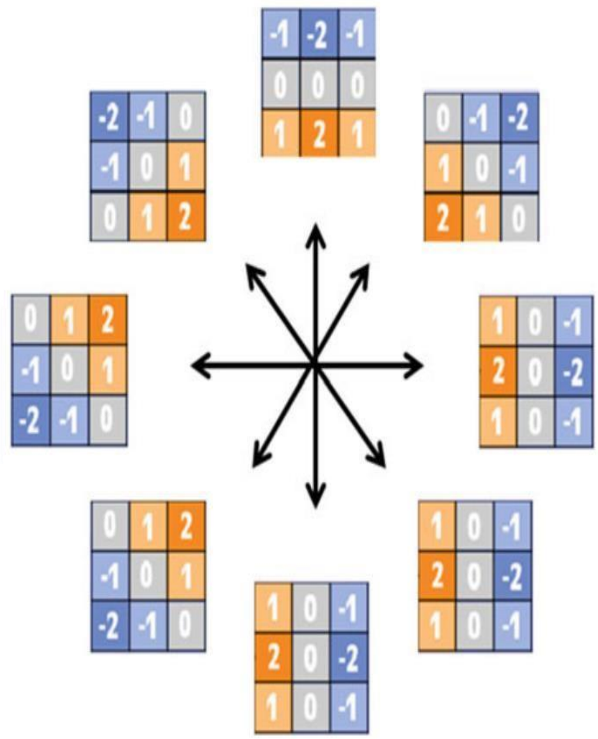


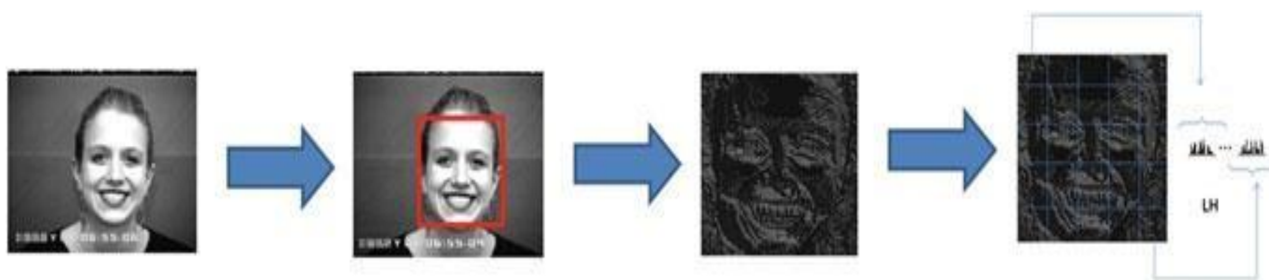
Figure 2.5. Robinson Mask

Encode prominent regions i.e., binary codes are the binary numbers of the directions that they represent. LDN is the concatenation of minimum and maximum directions.

Table 3: Direction and its equivalent binary number

Direction	Binary Number
East	000
North-East	001
North	010
North-West	011
West	100
South-West	101
South	110
South-East	111

Histogram Generation Phase: Each face is represented using LDN Histogram (LH). It contains fine to coarse information about an image, such as spots, edges, corners, and other local texture features. To extract location information, first divide the LDN image into small regions $\{R^1, R^2 \ldots R^N\}$. Then extract histogram H^i from each region R^i which represents minute details about local features of the face. We can divide the image into 3 × 3, 5 × 5, and 7 × 7. After obtaining a histogram from each block, concatenate these histograms which represent local features to form a global descriptor called LDN Histogram (LH). The entire process is shown in Figure 2.6



Input image | Preprocessing by cropping face region | LDN Code | LDN Histogram

Figure 2.6. Computation of LDN code

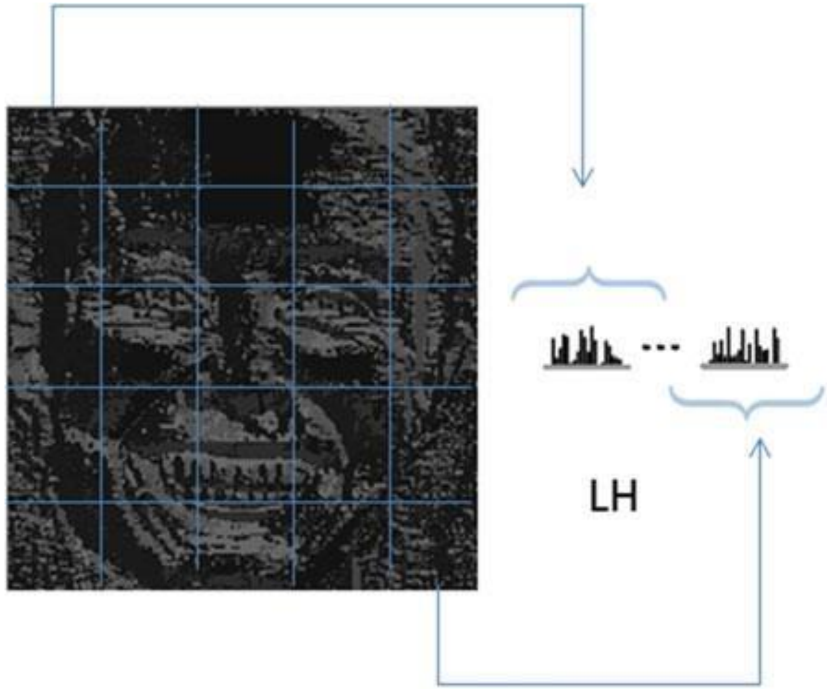


Figure 2.7. LDN Histogram

Dimensionality reduction Phase: Principal components analysis (PCA) searches for k n-dimensional orthogonal vectors that can best be used to represent the data, where $k \leq n$. It can be applied to ordered and unordered attributes, and can also handle sparse and skewed data. Data of more than two dimensions (multidimensional data) can be reduced to two dimensions in order to handle the problem. The dimensionality of LH can be reduced using PCA, i.e., the number of bins in the LDN histogram is reduced in such a way that the most prominent features are extracted.

Facial Expression Recognition Phase: The main objective is to compare an encoded feature vector of the test image with the feature vectors that are already trained. First, train the basic 6 expressions of an individual, and when a test image is given as input, it compares the data that are already trained and returns which expression corresponds to that image. This can be done with the help of the chi-square dissimilarity test.

The formula for *Chi-Square Test*:

$$X_c^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$

Where c = Degrees of freedom, O = Observed Value, E = Expected Value

2.2.3 RESULTS & DISCUSSIONS

The validation of this proposed system is done by using the images from the CK+ dataset which is available online. This system exhibited an average time of 8.32 s for expression recognition. This technology is profiled on a Windows laptop with a 2.00 GHZ processor. Table 4 shows the comparison study on the accuracy of our system for expression classification for Kirsch and Robinson Mask. Tested using 10 test images for each expression for a particular subject.

Table 4: The comparison study on the accuracy of our system for expression classification for Kirsch and Robinson Mask

Basic six expression	Compass mask used	
	Kirsch (%)	**Robinson (%)**
Surprise	90	80
Sad	100	100
Disgust	100	90
Anger	90	90
Happiness	90	80
Fear	100	90

The below Figure. 2.8 shows the histograms that we obtain while using Kirsch and Robinson Mask. For expression classification, using Kirsch Mask is more efficient because some of the distinct features are not accurately extracted while using Robinson Mask. The Figure 2.8 shows our pre-processed image obtained using Viola Jones Algorithm and its corresponding histogram.

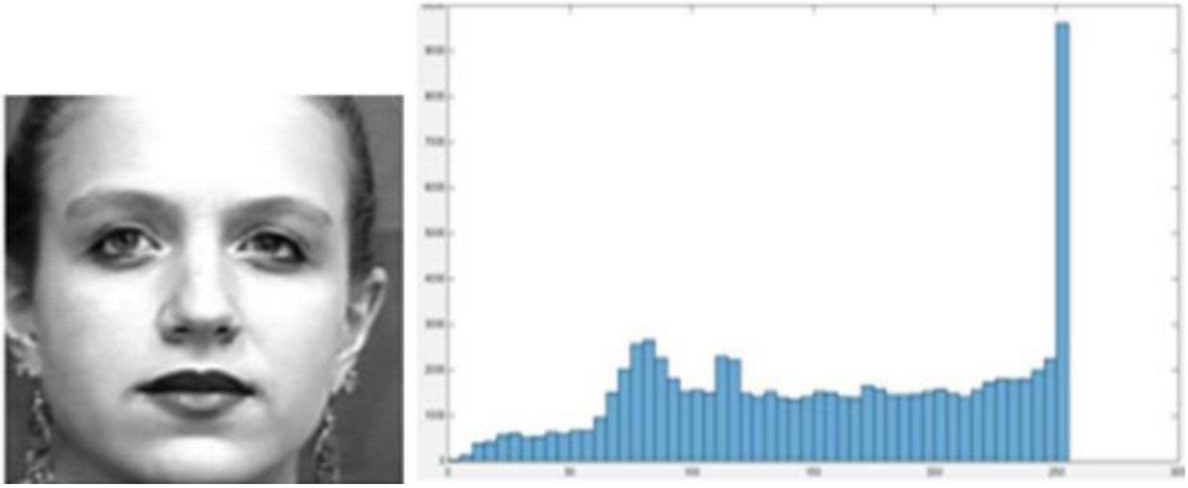


Figure 2.8. Pre-processed image and its corresponding histogram

The aim of the method is to extract facial relevant features. To accomplish the task, LDN code was generated using both the Kirsch mask and Robinson Mask. The result of the LDN code generated using the Kirsch mask is shown in the Figure 2.9. and the result of the LDN code generated using the Robinson mask is shown in the Figure 2.10.

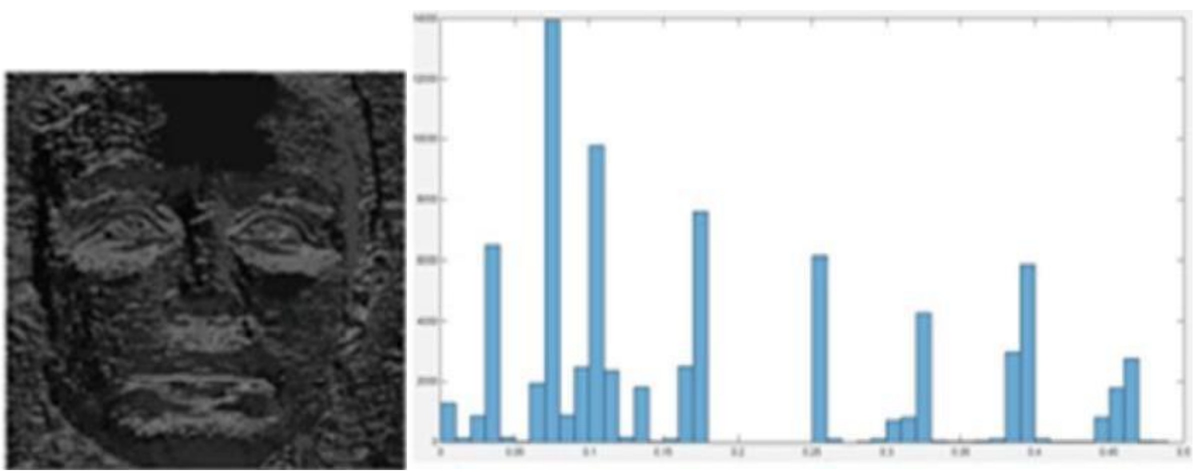


Figure 2.9. LDN code generated using Kirsch mask and its corresponding histogram.

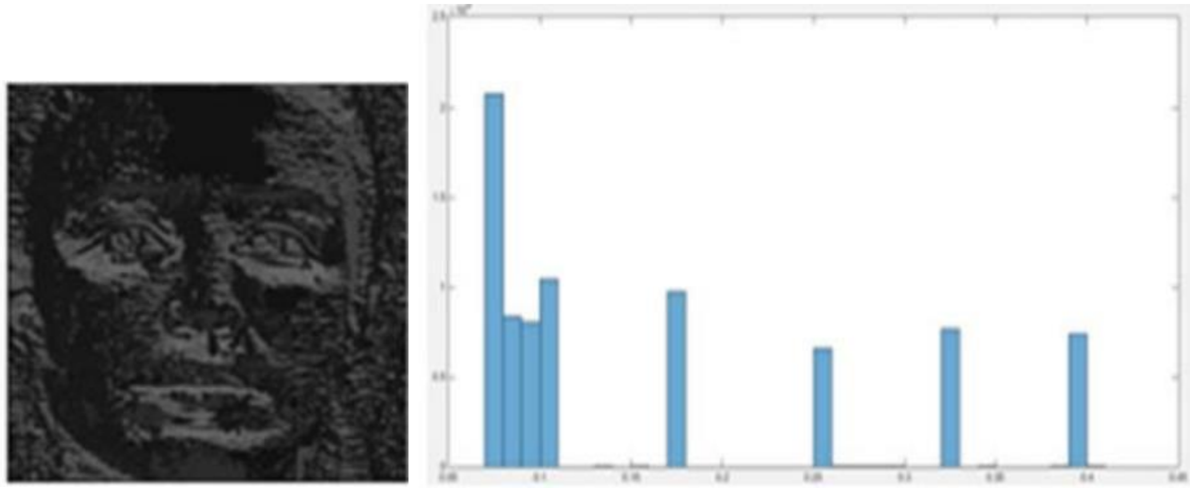


Figure 2.10. LDN code generated using Robinson mask and its corresponding histogram

LDN code generated using the Robinson mask are created. From their corresponding histograms, this method concluded that more relevant information is extracted while using the Kirsch mask compared to Robinson Mask.

2.3 FACIAL EXPRESSION RECOGNITION USING LOCAL DIRECTIONAL PATTERN (LDP)

2.3.1 INTRODUCTION

Local Directional Number (LDN) is a face descriptor, which evaluates the directional information of a local neighborhood to encode its structure. Mainly, LDN is a texture descriptor and it encodes the facial texture of structural information and intensity variations.

Local Binary Pattern (LBP), originally introduced for texture analysis, has been successfully applied as a local feature extraction method in facial expression recognition. Though LBP operator shows robustness to monotonic illumination change and computational efficient, it is sensitive in non-monotonic illumination variation and shows poor performance in presence of random noise. In this method, the analysis of the shortcomings of LBP features and present the superiority of the new facial feature Local Direction Pattern (LDP) for facial expression recognition is done. The performance of proposed LDP feature is evaluated with two machine learning methods first among them is Template matching and the other is Support Vector Machine (SVM) and it demonstrate that LDP features can be represented in lowdimensional feature space, while retaining discriminative facial information.

2.3.2 METHODOLOGY

In this method as Local Directional Pattern (LDP) is used and it is an eight-bit binary code assigned to each pixel of an input image. This pattern is calculated by comparing the relative edge response value of a pixel in different directions. Eight directional edge response were calculated using Kirsch masks in eight different orientations (M_0~M_7) centered on its own position as shown in Figure 2.11.

$$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

East M_0 North East M_1 North M_2 North West M_3

$$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

West M_4 South West M_5 South M_6 South East M_7

Figure 2.11. Kirsch edge response masks in eight directions

With eight directional masks, we obtain eight edge response value m0, m1, …, m7, each representing the edge significance in its respective direction. The response values are not equally important in all directions. The presence of corner or edge show high response values in particular direction. This method is interested in knowing the k most prominent directions in order to generate the LDP code. Hence, finding of the top k values |mj| and set them to 1. The other (8-k) bits of the 8- bit LDP pattern is set to 0.

$$C\ [\ f\,(x, y)] : = (c_i = 1) \text{ if } 0 \leq i \leq 7 \text{ and } m_i \geq \Psi$$

$$\text{Where } \Psi = \text{kth } (M);\ M = \{m0, m1, \ldots, m7\}$$

The Figure 2.12 demonstrates the generation of LDP code considering three most prominent edge responses.

85	32	26
53	50	10
60	38	45

Mask Index	M7	M6	M5	M4	M3	M2	M1	M0
Mask Value	161	97	161	537	313	97	-503	-393
Rank	6	7	5	1	4	8	2	3
Code Bit	0	0	0	1	0	0	1	1
LDP Code	19							

Figure 2.12. LDP Code generation

The Figure 2.13. shows an original image and the corresponding image after adding Gaussian white noise. The LDP code produces more stable pattern in presence of noise. For instance, Figure 2.13. shows an original image and the corresponding image after adding Gaussian white noise. After addition of noise, 5th bit of LBP changed from 1 to 0, thus LBP pattern changed from uniform to a non-uniform code. Since gradients are more stable than gray value, LDP pattern provides the same pattern value even presence of that noise and non-monotonic illumination changes.

85	32	26
53	50	10
60	38	45

LBP = 00111000

LDP = 00010011

(a)

81	29	32
38	58	15
65	43	47

LBP= 00101000

LDP = 00010011

(b)

Figure 2.13. Stability of LDP vs LBP (a) Original Image (b) Image with noise

2.3.2.1 FACIAL IMAGE REPRESENTATION

Facial image representation is carried out by an LDP histogram. For this reason, the first LDP-coded image is generated from the original face image by applying the LDP operator over the original face image to get an encoded image I_L. Use k=3 which generates $8C_k$ = 56 distinct values in the encoded image. So histogram H of this LDP labeled image $I_L(x,y)$ becomes a 56 bin histogram and can be defined by

$$H_i = \sum P\ \{I_L(x,y) = C_i\} \text{ here } C_i = i^{th} \text{ LDP pattern value, } i=0,1,2,3\ldots 8C_k$$

Here C_i = ith LDP pattern value, i=0, 1,….., ${}^{8}C_k$

$$
P(A) = \begin{cases} 1, & \text{If A is True} \\ 0, & \text{If A is False} \end{cases}
$$

LDP histogram contains fine detail information of an image, such as, edges, spots, corners, and other local texture features. But histogram computed over the whole face image encodes only the occurrences of the micro-patterns without any knowledge about their locations. In order to incorporate some degree of location information, we divide face images into small regions R0, R1,…, Rn and extracted the LDP histograms H Ri from each region Ri. These n LDP histograms are concatenated to get a spatially combined LDP histogram which plays the role of a global face feature for the given face image. Process can be visualized by Figure 2.14.

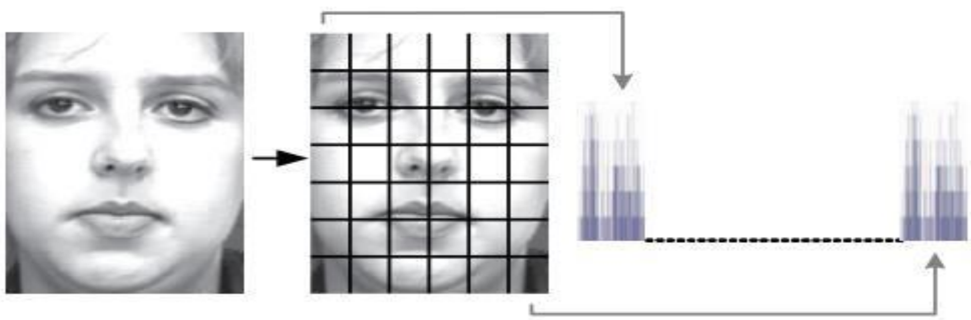


Figure 2.14. Facial image representation using spatially combined histogram

2.3.2.2 FACIAL EXPRESSION RECOGNITION

Person-independent facial expression recognition is performed using above mentioned LDP features. Machine learning techniques: such as template matching or support vector machines (SVM), are utilized for the recognition.

2.3.2.2.1 TEMPLATE MATCHING

Initially, this method adopted template matching for its simplicity to classify facial expressions. In the training step, the LDP histograms of expression images in each class were averaged to generate a template for that class. To match the test expression, the LDP histogram is compared with the template histogram with a weighted chi-square similarity measure.

$$\chi_w^2(SLH^1, SLH^2) = \sum_{i,j} w_i \frac{\left(SLH_{i,j}^1 - SLH_{i,j}^2\right)^2}{\left(SLH_{i,j}^1 + SLH_{i,j}^2\right)}$$

Here SLH1 and SLH2 are histograms of template expression and histogram of test expression respectively, the index ***i*** refer to the region number, ***j*** indicates bin number of that region and ***wi*** is the associated weight of region i. Here usage of weighted version of chi-square similarity is done because some local facial regions contain more useful information for expression classification than others. For example, facial features contributing to facial expressions mainly lie in regions such as eye and mouth regions. Therefore, a weight can be set to each of those sub-regions based on its importance. The weights adopted in this method are shown in the below Figure 2.15.

Figure 2.15. A facial image divided in 7x6 sub-regions. The eights assigned for the weighted χ^2 dissimilarity measure. Black indicates weight of 0.0, dark gray indicates 0.5, light gray indicates 1.0 and white indicates 1.5.

2.3.2.2.2 SUPPORT VECTOR MACHINE

SVM is a popular supervised machine learning technique which performs an implicit mapping of data into a higher dimensional feature space. After that it finds a linear separating hyperplane with maximal margin to separate data from this higher dimensional space.

Given a training set of labelled examples T = { (x, y), i = 1…M } where $x_i \in Rn$ and $y_i \in$ {-1,1}, the new test data will be classified by equation

$$f(x) = \text{sign} \sum_{i=1}^{M} \alpha i y i K(xi, x) + b$$

where α_i are Lagrange multipliers of dual optimization problem, b is a bias and *K(.,.)* is a kernel function SVM allows domain-specific selection of the kernel function. Though new kernels are being proposed, the most frequently used kernel functions are the linear, polynomial, and Radial Basis Function (RBF) kernels.

SVM makes binary decisions and multi-class classification can be achieved by adopting the one-against-one or one-against-all techniques. In this method adoption of one-against-one technique which constructs k(k-1)/2 classifiers where each one is trained on data from the two classes. Grid-search is carried out on the hyperparameters in the 10-fold cross-validation for selecting the parameters, as suggested in. The parameter setting producing best cross-validation accuracy was picked.

2.3.3 RESULTS AND DISCUSSIONS

The performance of this method was evaluated with the image dataset from Cohn- Kanade Facial Expression Database. This method evaluates based on the images that were selected from the database for basic expression recognition. These sequences came from 96 subjects, with 1 to 6 expressions per subject. For the seven class classifications one neutral frame and three most expressive frames while for six class classifications only three most expressive frames are taken from each expression sequence. These selected image frames are then cropped into 110 x 150 pixels using position of two eyes of the expresser. The average recognition rate is shown in Table 5 which exhibits superiority of LDP feature in person-independent expression recognition over other existing facial features.

Table 5. Recognition performance using different method

Methods (Feature + Classifier)	7-class Recognition (%)	6-class Recognition (%)
LBP + Template Matching	79.1 ± 4.6	84.5 ± 5.2
LBP + SVM	88.9 ± 3.5	92.6 ± 2.9
Gabor + SVM	86.8 ± 3.6	89.8 ± 3.1
LDP+ Template Matching	86.9 ± 2.8	89.2 ± 2.5
LDP + SVM	93.4 ± 1.1	96.4 ± 0.9

In case of six class recognition, LDP feature achieves excellent accuracy with SVM. The confusion matrix of 6-class recognition is shown in Table 6 which exhibits individual classification accuracy of each expression is over 95%.

Table 6. Confusion matrix of 6-class facial expression

Expression	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sad (%)	Surprise (%)
Anger	95.56	2.48	0.0	0.0	1.48	0.48
Disgust	0.0	96.5	3.5	0.0	0.0	0.0
Fear	1.51	0.0	95.68	2.51	0.0	0.0
Joy	0.0	0.0	0.0	98.0	0.0	2.0
Sad	0.53	1.53	0.0	0.0	97.94	0.0
Surprise	0.0	0.0	0.0	3.0	0.0	97.0

The confusion matrix of 7-class recognition is also shown in Table 7, which ensures that other than anger and neutral expression all five expressions can be classified with reasonable higher accuracy. Only anger and neutral expressions are sometimes made confusion with each other.

Table 7. Confusion matrix of 7- class facial expression

Expression	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sad (%)	Surprise (%)	Neutral (%)
Anger	**86.9**	0.9	0.9	0.0	0.0	0.9	10.4
Disgust	2.0	**94.2**	0.0	0.0	0.0	0.0	3.8
Fear	1.5	0.0	**94.4**	0.0	0.0	0.0	4.1
Joy	0.0	0.0	0.7	**98.9**	0.0	0.0	0.4
Sad	1.1	0.5	0.0	0.0	**92.6**	0.0	5.8
Surprise	0.0	0.0	0.0	0.0	0.0	**99.0**	1.0
Neutral	5.9	1.2	0.7	0.0	2.7	0.2	**89.3**

The performance of the proposed LDP feature in low-resolution images are also investigated. For this purpose, sub-sampling of original image into three different image resolutions (55x75, 36x48 and 27x37) and have evaluated the performance of the recognition Table 8 shows that even in the lower resolution image our proposed facial feature descriptor (LDP) can effectively classify the expression with reasonably higher accuracy.

Table 8. Recognition performance in low-resolution images

Methods (Feature + Classifier)	**Face Image Resolutions**		
	55 x 75	36 x 48	27 x 37
LBP + SVM	89.9 ± 3.1	87.3 ± 3.4	84.3 ± 4.1
Gabor + SVM	89.2 ± 3.0	86.4 ± 3.3	83.0 ± 4.3
LDP + SVM	95.5 ± 1.6	93.1 ± 2.2	90.6 ± 2.7

2.4 SUMMARY

This chapter discussed various existing methods for facial expression recognition. All the methods provide feature extraction and then classification is done based on any classification algorithm such as Template matching, Support Vector Machine (SVM), etc.

One of the methods is **Local Binary Pattern (LBP)** which has advantages like its invariance to monotonic gray-level changes and computational efficiency thus making it suitable for demanding image analysis tasks as it is a local descriptor and it is robust to illumination changes and it serves with the limitations like usage of larger local regions increases its robustness to errors.

The other method used for facial expression recognition is **Facial Expression Recognition Using PCA and Texture-Based LDN Descriptor** this method compares the usage of two masks namely the **Kirsch Mask** and **Robinson Mask** for facial expression recognition and at last, this method concludes that the usage of Kirsch mask is more efficient than the Robinson mask. It has advantages like as it uses LDN so it helps to distinguish among similar structural patterns having different intensity transitions besides this method exhibits an average time of 8.23s for expression recognition and limitations like some of the irrelevant features which are not related to facial expression are also extracted which reduces the efficiency of the overall proposed method.

The next method for facial expression recognition discussed in this chapter is **Facial Expression Recognition Using Local Directional Pattern (LDP)** it is more efficient compared to other existing methods because in this method all the shortcomings of LBP were analyzed first and eradicated in this method apart from that using LDP as a facial feature leads to the recognition of facial expression in low resolution images can also be achieved the classification methods used in this method were Template Matching and Support Vector machine which reveals that using SVM+LDP gives more accuracy when compared to SVM with another Feature. The advantage of using this LDP is its insensitive nature to noise and non-monotonic illumination variations. Hence the facial expression recognition system using this LDP feature can classify different expressions with higher accuracy. The limitation of this method is that as this method considers the smooth regions for evaluation of facial expression recognition which somehow reduces the accuracy of the method.

CHAPTER 3

LOCAL DIRECTIONAL TERNARY PATTERN FOR FACIAL EXPRESSION RECOGNITION

3.1 INTRODUCTION

Facial expressions can be represented by appearance changes on the face. Describing them exactly is the key issue in facial expression recognition for detecting emotions. There are two main approaches to describe facial images one of them is geometric-feature-based and the other is appearance-feature-based methods. The first method represents the facial image by encoding location relations of main facial components, like eyes, nose, mouth, etc. However, the recognition performance relies on the exact locations of key facial components, which are difficult to detect under appearance changes on the face according to facial expressions.

The appearance-feature-based methods can avoid this problem innately. They represent the facial image by using image filters which are applied on the whole face (holistic) or specific - face regions (local) to extract the appearance variations of facial images. Holistic methods, can accommodate local variations that occur by expression changes. Methods that extract edge-based local features and histogram representation proved successful in facial expression recognition as emotion-related facial features have prominent gradient magnitudes. Thus, their histogram representation is robust to small location and code errors.

Extracting edge-based local features in the smooth regions of the face image makes unstable patterns which are sensitive to noise and contribute negatively to the classification result. Spatial information of the face features plays an important role in the expression recognition, but histogram representation is inefficient to preserve spatial information. To increase spatial information in the histogram representation, the number of uniform regions should be increased. The performance degradation is more significant at the smooth region where no prominent and stable edge patterns exist.

A new face descriptor called Local Directional Ternary Pattern (LDTP) for facial expression recognition. Motivated by the high edge responses in the boundaries of the emotion-related facial features, it extracts edge directional patterns in a face image, while avoiding to generate ones from smooth regions by using the magnitude of the edge response. It extracts two main edge directions as directional patterns at each local pixel and utilize them to extract a local feature only if the edge response is higher than a threshold determined from experiments. To encode the validation and sign information of an edge direction, add a ternary pattern to each directional pattern. A way to select active edge patterns which have significant accumulation for histogram and positional variation among facial expressions. Based on selective edge patterns, a new coding scheme that increases spatial information while suppresses the sampling error, which results in better classification performance in overall recognition. The proposed coding scheme assigns positional bits only to the active edge patterns with significant accumulation to get the effect of using finer grid for the selected codes. It increases the overall performance by applying the finer grid to the active patterns.

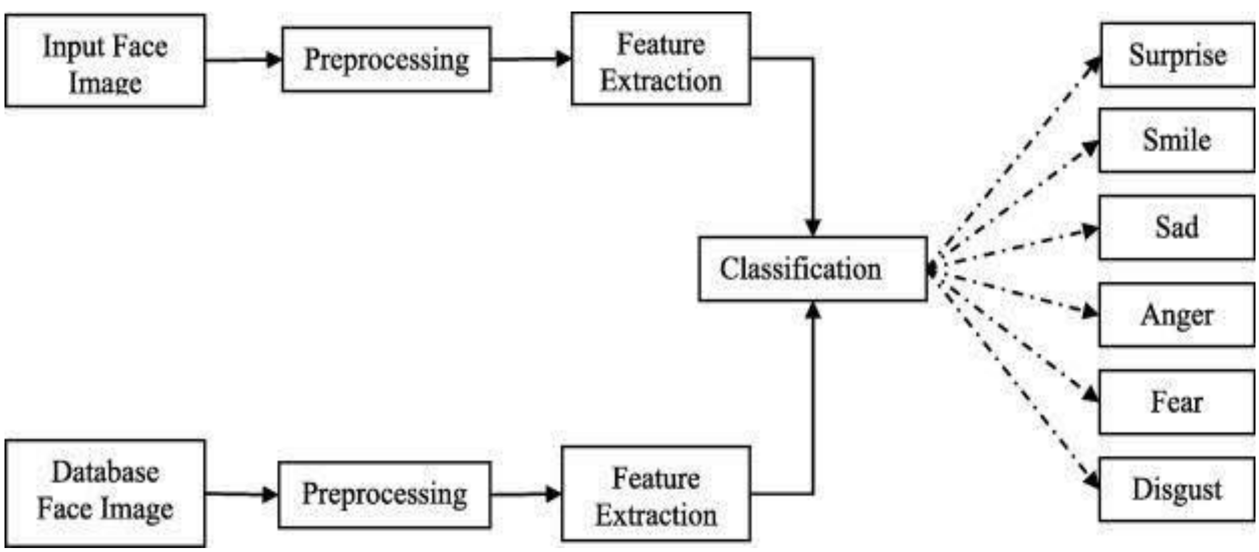


Figure 3.1. Block diagram for Facial Expression Recognition

3.2 METHODOLOGY

In this method feature extraction is done from already preprocessed images which are available in the CK+ dataset.

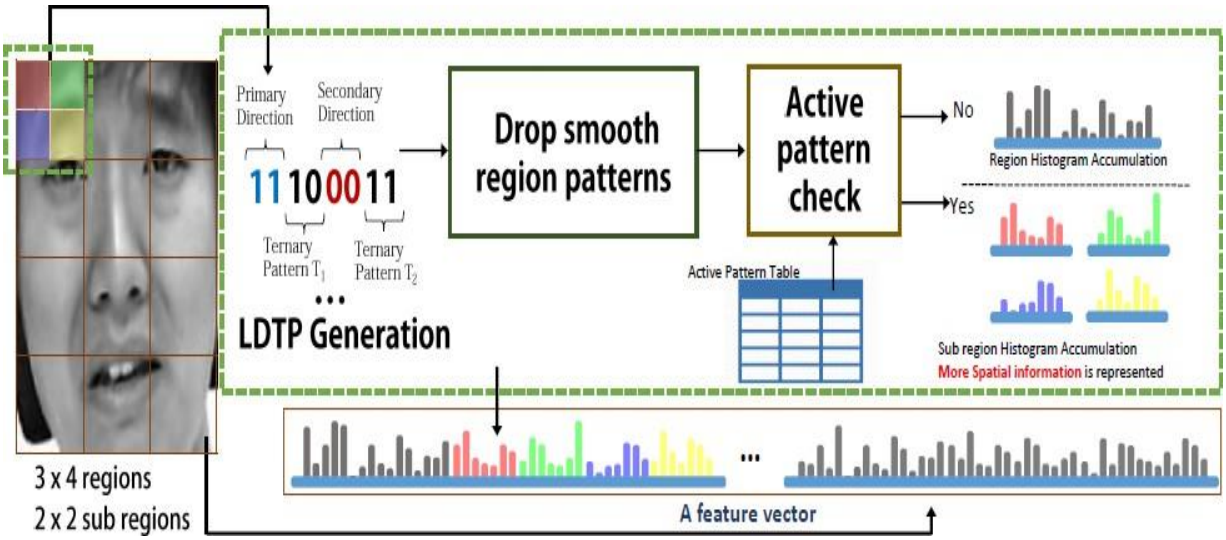


Figure 3.2. overall procedure of the proposed method

The Figure 3.2 shows the overall procedure of the proposed method. For each face image, first calculate LDTP codes by using edge response and drop codes from smooth areas. Then, for each region, add positional bits to active patterns and generated codes are accumulated into a histogram. Lastly, all histograms are concatenated into a feature vector for facial expression recognition.

3.2.1 CODING SCHEME (FEATURE EXTRACTION)

Local Directional Ternary Pattern (LDTP) is an eight-bit pattern code assigned to each pixel of an input face image. In expression recognition, the shape of the facial features that change according to expressions is more influential than whole-face textures used in face recognition, and the boundaries of the facial features have high edge magnitudes. Therefore, use Robinson compass masks as an edge operator to calculate edge responses efficiently, and take two main directions at each pixel to represent local edge shapes. Robinson Compass Masks are shown in Figure 3.3.

$$
\underset{M_0}{\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}}
\underset{M_1}{\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}}
\underset{M_2}{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}}
\underset{M_3}{\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}}
$$

$$
\underset{M_4}{\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}}
\underset{M_5}{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}}
\underset{M_6}{\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}}
\underset{M_7}{\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}}
$$

Figure 3.3. Robinson compass masks used for computing the directions on LDTP.

The properties of LDTP are Gradient direction is used for the superior representation of shapes of the emotion-related facial features, Robinson compass mask is efficient due to its symmetry, The ternary pattern encodes edge-sign information and differentiates between edge and smooth (non-edge) regions (thus, solving the weakness of edge patterns in smooth areas). As Robinson compass masks are symmetric and generate the same magnitude response with different signs in opposite directions. Therefore, we can use only four masks from M_0 to M_3 to find the principal directions, which can reduce calculation time. So, in order to form a ternary pattern, encode the symmetrical representation by using four directional codes and the sign information. Assign 2 bits to encode the primary directional number, and 2 bits for the secondary one and each directional number has 2 bits for each ternary pattern.

Robinson's compass mask is applied over the entire image producing a set of response magnitudes correlated with the four directions by using the formula:

$$Ri = Mi * I, \; 0 <= i <= 3;$$

Where I is the original image, Mi is the ith Robinson compass mask, and Ri is the ith response image. The next step After getting the response images (R_0 to R_3) search for the jth maximum absolute value Dj of the four Robinson compass mask's responses by using the given formula

$$D_j(x, y) = \underset{i}{\overset{j}{\operatorname{argmax}}}\{\, |R_i(x, y)| : 0 \le i \le 3\}$$

Where $\arg\max_i^j$ is an operator that returns the index i of the jth maximum value in the set. The values of j are 1 and 2 because there are two directions.

Next find the ternary pattern which is used to determine whether the direction is in an edge or in a smooth area that can be computed by using formula:

$$T_j(x, y) = \begin{cases} 2 \ \text{if } R_i(x, y) < -\sigma, \\ 1 \ \text{if } R_i(x, y) > \sigma, \\ 0 \ \text{if } -\sigma \leq R_i(x, y) \leq \sigma \end{cases}$$

Where T_j is the ternary pattern of the magnitude of the j^{th} direction at position (x,y), R_i (x,y) is the edge response of ith direction at position (x,y), i = D_j(x,y) is the jth principal direction at position (x,y), and σ is a threshold value which is adaptively changes from image to image by considering median of the response image R_i. The threshold divides the data only into three sections, upper, lower, and in between. Upper means a strong positive and lower means a negative edge response, whereas in between means a weak edge response. With this differentiation, it is possible to separate and keep the directional information from the edges response and ignore the directional information of the smooth areas.

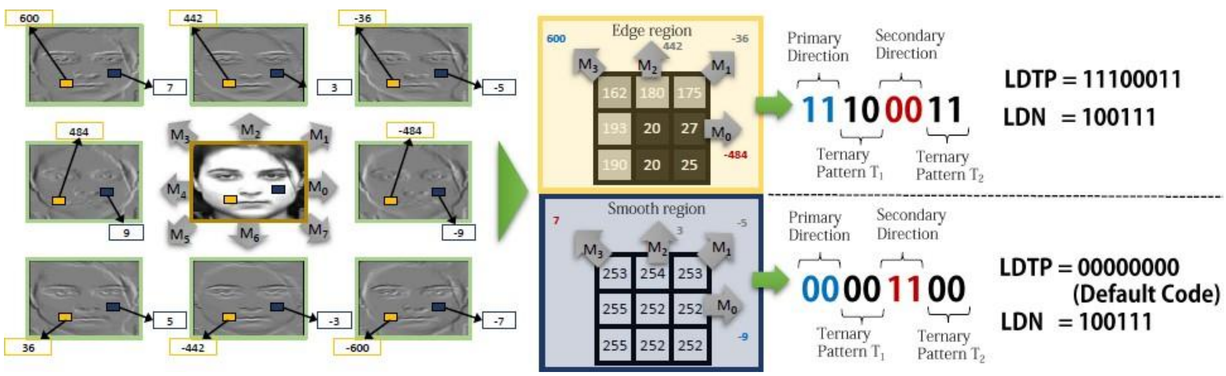


Figure 3.4. Mask Convolution Results and LDTP Code Computation

The Figure 3.4 shows the LDTP code computation. Calculation of the edge response by Robinson compass masks with the original image, and select the primary and secondary direction to encode the shape of facial features. The above figure is an example of two 3X3 image patches corresponding to the edge responses. LDTP can detect smooth regions by making a different code while other gradient-based patterns cannot (like LDN), as they produce the same code for remarkably different textures.

Last step in this feature extraction process is to compute the LDTP codes. If the ternary pattern from the first direction is 0, that means that the pixel (x, y) exists on a smooth area and therefore an empty code (0) is generated which later is being ignored. If the ternary pattern from the second direction is 0, only the information of the first direction is meaningful and the knowledge of the second direction can be discarded. The code is created by concatenating the binary form of the two principal directions and the two ternary patterns. This concatenation can be represented by the following operation.

$$\mathrm{LDTP}\,(x, y) = 2^6\,\mathrm{D}_1\,(x, y) + 2^4\,\mathrm{T}_1\,(x, y) + 2^2\,\mathrm{D}_2\,(x, y) + \mathrm{T}_2\,(x, y)$$

Where LDTP *(x, y)* is the code for each pixel *(x, y)* in the face image, D1 and D2 are the direction number of the primary and secondary directions, and T1 and T2 are the first and second ternary patterns of the two directions, respectively.

After generation of LDTP codes for a given facial image convert the LDTP codes into its equivalent histogram which gives the facial features. Histogram based on the LDTP codes that are generated for a facial image is shown in Figure 3.5.

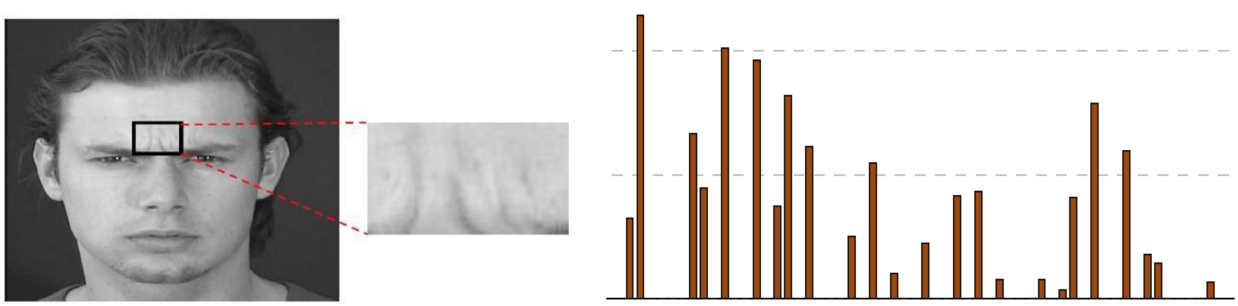


Image patch 1 LDTP Histogram of patch 1

Figure 3.5. LDTP code generated for an image patch

3.2.2 CLASSIFICATION

After the completion of feature extraction classification is needed to be done. The classification based on the obtained facial features. In this method K-Nearest Neighbour (K-NN) is used as a classifier for the classification of new data. If f1, f2,.. fk are the feature vectors and (*x*, *y)* be the data point then by using the k-NN algorithm the distance from the new data to all other data that is already classified is computed using the formula.

$$\sqrt{\sum_{i=1}^{k}(xi - yi)^2}$$

Labelling to the new data point is done given based on the obtained nearest neighbour. All the image data used for this classification is taken from the CK+ dataset.

CHAPTER 4

4. PERFORMANCE EVALUATION

4.1 EXPERIMENTAL SETUP

Experiments of facial expression recognition are done to validate the efficiency of the proposed method. As a part of pre-processing the images are cropped into 110 X 150 pixels and normalized according to the position of the eyes and mouth. Training to the proposed model is done by using the images from the famous database CK+. The K-Nearest Neighbor is used as a classifier to classify the input image and assign a label to it based on the facial expression. The model is evaluated using different performance metrics.

4.2 PERFORMANCE METRICS

The evaluation of a model is performed with the help of confusion matrix. Totally, four outcomes are generated by confusion matrix, namely TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative). The measures used for the calculation of accuracy, sensitivity, specificity, precision, F-measure are as follows:

- Accuracy = (TP+TN) / (TP+TN+FN+FP)
- Sensitivity (Recall) = TP / (TP+FN)
- Specificity = TN / (TN+FP)
- Precision = TP / (TP+FP)
- F-Measure = 2TP / (2TP+FP+FN)

The evaluation of a model is performed with the help of confusion matrix.

1. **TruePositive (TP):** The prediction is positive and the actual value is positive.
2. **True Negative (TN):** The prediction is negative and the actual value is negative.
3. **False Positive (FP):** Prediction is negative but the actual value is positive.
4. **False Negative (FN):** Prediction is but positive the actual value is negative.

4.3 RESULTS

From the dataset, 30 random images of different expressions were taken for evaluating the proposed model. For performance measurements different measures were used which were based on the performance metrics.

Accuracy: Accuracy is computed as ‘‘the total number of correct predictions, True Positive (TP)+True Negative (TN) divided by the total number of a dataset Positive (P)+Negative (N)’’.

Accuracy= (TP+TN) / (TP+TN+FP+FN).

Precision: Precision is computed as ‘‘the number of correct positive predictions (TP) divided by the total number of positive predictions (TP + FP)’’. Precision is also known as a positive predictive value.

Precision= (TP) / (TP+FP).

Recall: Recall is computed as ‘‘the number of correct positive predictions (TP) divided by the total number of positives (P)’’. Recall is also known as the true positive rate or sensitivity.

Recall = TP /(TP+FP).

Confusion Matrix

$$\begin{bmatrix} 9 & 3 \\ 1 & 17 \end{bmatrix}$$

[[9 3]

[1 17]]

Results obtained based on the above generated confusion matrix are given below

Specificity= 85.71428571428571

Sensitivity= 90.0

Precision= 75.0

F-Measure= 81.81818181818183

Accuracy: 86.666666666667 %

4.3.1 SUBJECTIVE RESULTS

INPUT IMAGE **OUTPUT**

Detected Emotion is: Happy

Detected Emotion is: Sad

Detected Emotion is: Angry

Detected Emotion is: Surprise

4.4 SUMMARY

In this chapter a new facial expression descriptor Local Directional Ternary Pattern (LDTP) is introduced and moreover how it is robust than the existing methods is discussed like removing the smooth regions in the facial image for evaluation considering that smooth regions do not contribute anything to the facial images. Then how the classification of the facial features based on the K-Nearest Neighbor (K-NN) is discussed. After the classification the evaluation of the proposed method is done using different performance metrics.

The main advantage of this proposed method is that not consideration of smooth regions for extracting the feature and LDTP efficiently encodes information of emotion-related features (i.e., eyes, eyebrows, upper nose, and mouth) by using the directional information and ternary pattern in order to take advantage of the robustness of edge patterns in the edge region while overcoming weaknesses of edge-based methods in smooth regions.

The directional information is suitable to describe shapes of emotion-related facial features, which makes LDTP a more discriminable and robust pattern than existing methods for facial expression recognition. And the use of ternary pattern makes the proposed LDTP produce more reliable and stable codes than existing edge based methods since it removes uncertainty of directional pattern generated in smooth region.

CHAPTER 5

CONCLUSION

A new local pattern, Local Directional Ternary Pattern (LDTP), that efficiently encodes shapes of emotion-related features by using the directional information has been proposed. For robust encoding LDTP incorporates ternary patterns that allow it to distinguish directional patterns on edge or smooth regions in which arbitrary, meaningless, and noise sensitive patterns are generated. The KNN classifier is used to test the performance of facial expression. The efficiency is analysed using Precision and recall. The overall accuracy 86.667% is generated by using the Local Directional Ternary Pattern (LDTP) as a facial descriptor.

REFERENCES

[1] P. Weinberg and W. Gottwald, “Impulsive consumer buying as a result of emotions,” Journal of Business research, vol. 10, no. 1, pp. 43–57,1982.

[2] D. Consoli, “Emotions that influence purchase decisions and their electronic processing,” Annales Universitatis Apulensis: Series Oeconomica, vol. 11, no. 2, p. 996, 2009.

[3] Y.-L. Tian, T. Kanade, and J. F. Cohn, “Facial expression analysis,” in Handbook of face recognition. Springer, 2005, pp. 247–275.

[4] H. Hong, H. Neven, and C. von der Malsburg, “Online facial expression recognition based on personalized galleries,” in Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, apr 1998, pp. 354 –359.

[5] N. Bourbakis and P. Kakumanu, “Skin-based face detection-extraction and recognition of facial expressions,” in Applied Pattern Recognition. Springer, 2008, pp. 3–27.

[6] N. Bourbakis, A. Esposito, and D. Kavraki, “Extracting and associating meta-features for understanding peoples emotional behaviour: Face and speech,” Cognitive Computation, vol. 3, no. 3, pp. 436–448, 2011.

[7] P. Kakumanu and N. Bourbakis, “A local-global graph approach for facial expression recognition,” in Tools with Artificial Intelligence, 2006. ICTAI’06. 18th IEEE International Conference on. IEEE, 2006, pp.685–692.

[8] X. Xie and K.-M. Lam, “Facial expression recognition based on shape and texture,” Pattern Recognition, vol. 42, no. 5, pp. 1003–1011, 2009.

[9] E. Revealed, “Recognizing faces and feelings to improve communication and emotional life,” Henry Holt & Co, 2003.

[10] M. Pantic and M. S. Bartlett, Machine analysis of facial expressions. I-Tech Educat and Publishing, 2007.

[11] M. Turk and A. Pentland, “Eigenfaces for recognition,” Journal of cognitive neuroscience,vol. 3, no. 1, pp. 71–86, 1991.

[12] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, “Eigenfaces vs fisherfaces: Recognition using class specific linear projection,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, no. 7, pp. 711–720, 1997.

[13] K. Etemad and R. Chellappa, “Discriminant analysis for recognition of human face images,” JOSA A, vol. 14, no. 8, pp. 1724–1733, 1997.

[14] Z. Nenadic, “Information discriminant analysis: Feature extraction with an information-theoretic objective,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, no. 8, pp. 1394–1407, 2007.

[15] J. Yang, D. Zhang, A. F. Frangi, and J.-y. Yang, “Two-dimensional pca: a new approach to appearance-based face representation and recognition,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, no. 1, pp. 131–137, 2004.

[16] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 12, pp. 2037– 2041, 2006.

[17] C. F. Shan, S. G. Gong, and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” Image and Vision Computing, vol. 27, no. 6, pp. 803 –816, 2009.

[18] S. Yang and B. Bhanu, “Facial expression recognition using emotion avatar image,” in Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. IEEE, 2011, pp.866–871.

[19] T. Jabid, M. H. Kabir, and O. Chae, “Robust facial expression recognition based on local directional pattern,” ETRI journal, vol. 32, no. 5, pp. 784–794, 2010.

[20] Ramirez Rivera, R. Castillo, and O. Chae, “Local directional number pattern for face analysis: Face and expression recognition,” Image Processing, IEEE Transactions on, vol. 22, no. 5, pp. 1740–1752, 2013.

[21] Ramirez Rivera and O. Chae, “Spatiotemporal directional number transitional graph for dynamic texture recognition,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. PP, no. 99, pp. 1–1,2015.

[22] M. Dahmane and J. Meunier, “Emotion recognition using dynamic gridbased hog features,” in Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. IEEE, 2011, pp. 884–888.

[23] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," Image Processing, IEEE Transactions on, vol. 19, no. 6, pp. 1635–1650, 2010.

[24] J. Ylioinas, A. Hadid, Y. Guo, and M. Pietik¨ainen, "Efficient image appearance description using dense sampling based local binary patterns," in Computer Vision–ACCV 2012. Springer, 2013, pp. 375–388.

[25] G. S. Robinson, "Edge detection by compass gradient masks," Computer Graphics and Image Processing, vol. 6, no. 5, pp. 492–501, 1977.

[26] O. Lahdenoja, M. Laiho, and A. Paasio, "Reducing the feature vector length in local binary pattern based face recognition," in Image Processing, 2005. ICIP 2005. IEEE International Conference on, vol. 2. IEEE, 2005, pp. II–914.

[27] T. Gritti, C. Shan, V. Jeanne, and R. Braspenning, "Local features based facial expression recognition with face registration errors," in Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on. IEEE, 2008, pp. 1–8.

[28] Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," Image Processing, IEEE Transactions on, vol. 19, no. 2, pp. 533–544, 2010.

[29] G. Lowe, "Distinctive image features from scale-invariant keypoints "International journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.

[30] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on. IEEE, 2010, pp. 94–101.

[31] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models " IEEE Transactions on pattern analysis and machine intelligence, vol. 23, no. 6, pp. 681–685, 2001.

[32] Matthews and S. Baker, "Active appearance models revisited," International Journal of Computer Vision, vol. 60, no. 2, pp. 135–164, 2004.

[33] W. V. Friesen and P. Ekman, “Emfacs-7: Emotional facial action coding system,” Unpublished manuscript, University of California at San Francisco, vol. 2, p. 36, 1983.

[34] P. Ekman, E. Rosenberg, and J. Hager, “Facial action coding system affect interpretation dictionary (facsaid),” 1998.

[35] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, “Coding facial expressions with gabor wavelets,” in Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, apr 1998, pp. 200 –205.

[36] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, “Web-based database for facial expression analysis,” in Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on. IEEE, 2005, pp. 5–pp.

[37] M. Valstar and M. Pantic, “Induced disgust, happiness and surprise: an addition to the mmi facial expression database,” in Proc. Intl Conf. Language Resources and Evaluation, Wshop on EMOTION, 2010, pp.65–70.

[38] T. Sim, S. Baker, and M. Bsat, “The cmu pose, illumination, and expression (pie) database,” in Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on. IEEE, 2002, pp. 46–51.

[39] “The cmu pose, illumination, and expression database,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 25, no. 12, pp. 1615–1618, 2003.

[40] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, “The first facial expression recognition and analysis challenge,” in Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. IEEE, 2011, pp. 921–926.

[41] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, “A 3d facial expression database for facial behavior research,” in Automatic face and gesture recognition, 2006. FGR 2006. 7th international conference on IEEE, 2006, pp. 211–216.

[42] M. J. Lyons, J. Budynek, and S. Akamatsu, “Automatic classification of single facial images,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 12, pp. 1357–1362, 1999.